

Week 2

# Area Plots

---

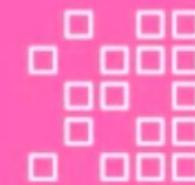
© IBM Corporation. All rights reserved.

# What you will learn

---



Describe what is an  
area plot



Explain how to create  
an area plot using  
Matplotlib

# Area plot

---

- Also known as area chart or graph
- Displays magnitude and proportion of multiple variables
- Represents time or another ordered dimension
- Is based on the line plot

# Dataset: Recap

---

	Type	Coverage	OdName	AREA	AreaName	REG	RegName	DEV	DevName	1980	...	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
0	Immigrants	Foreigners	Afghanistan	935	Asia	5501	Southern Asia	902	Developing regions	16	...	2978	3436	3009	2652	2111	1746	1758	2203	2635	2004
1	Immigrants	Foreigners	Albania	908	Europe	925	Southern Europe	901	Developed regions	1	...	1450	1223	856	702	560	716	561	539	620	603
2	Immigrants	Foreigners	Algeria	903	Africa	912	Northern Africa	902	Developing regions	80	...	3616	3626	4807	3623	4005	5393	4752	4325	3774	4331
3	Immigrants	Foreigners	American Samoa	909	Oceania	957	Polynesia	902	Developing regions	0	...	0	0	1	0	0	0	0	0	0	0
4	Immigrants	Foreigners	Andorra	908	Europe	925	Southern Europe	901	Developed regions	0	...	0	0	1	1	0	0	0	0	1	1

# Dataset: Processed

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Country																					
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	6
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

df\_canada

# Generating area plots

```
df_canada.sort_values(['Total'], ascending = False, axis = 0, inplace = True)
```

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Country																					
India	Asia	Southern Asia	Developing regions	8880	8670	8147	7338	5704	4211	7150	...	36210	33848	28742	28261	29456	34235	27509	30933	33087	691904
China	Asia	Eastern Asia	Developing regions	5123	6682	3308	1863	1527	1816	1960	...	42584	33518	27642	30037	29622	30391	28502	33024	34129	659962
United Kingdom of Great Britain and Northern Ireland	Europe	Northern Europe	Developed regions	22045	24796	20620	10015	10170	9564	9470	...	7258	7140	8216	8979	8876	8724	6204	6195	5827	551500
Philippines	Asia	South-Eastern Asia	Developing regions	6051	5921	5249	4562	3801	3150	4166	...	18139	18400	19837	24887	28573	38617	36765	34315	29544	511391
Pakistan	Asia	Southern Asia	Developing regions	978	972	1201	900	668	514	691	...	14314	13127	10124	8994	7217	6811	7468	11227	12603	241600

India

China

United Kingdom

Philippines

Pakistan

# Generating area plots

---

Country	India	China	United Kingdom of Great Britain and Northern Ireland	Philippines	Pakistan
1980	8880	5123	22045	6051	978
1981	8670	6682	24796	5921	972
1982	8147	3308	20620	5249	1201
1983	7338	1863	10015	4562	900
1984	5704	1527	10170	3801	668

# Generating area plots

```
years = list(map(str, range(1980, 2014)))

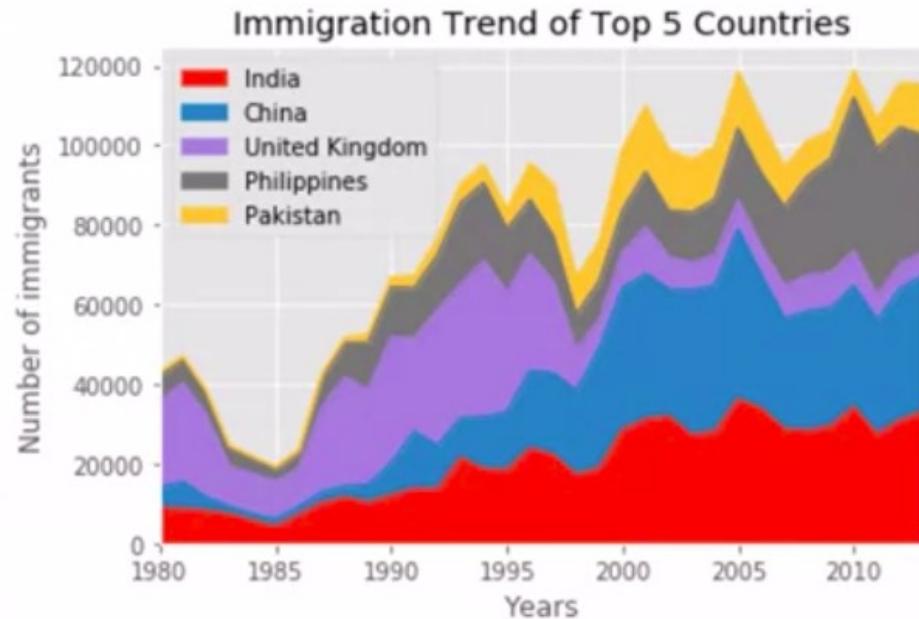
df_canada.sort_values(['Total'], ascending = False, axis = 0, inplace = True)

df_top5 = df_canada.head()
df_top5 = df_top5[years].transpose()
```

Country	India	China	United Kingdom of Great Britain and Northern Ireland	Philippines	Pakistan
1980	8880	5123	22045	6051	978
1981	8670	6682	24796	5921	972
1982	8147	3308	20620	5249	1201
1983	7338	1863	10015	4562	900
1984	5704	1527	10170	3801	668

# Area plots

```
import matplotlib as mpl  
import matplotlib.pyplot as plt  
  
df_top5.plot(kind='area')  
  
plt.title('Immigration trend of top 5 countries')  
plt.ylabel('Number of immigrants')  
plt.xlabel('Years')  
  
plt.show()
```



# Area plot: Uses

---

- Tracking stock market performance
- Visualizing population demographics
- Displaying the distribution of resources

# Recap

---

In this video, you learned that:

- An area plot depicts cumulated totals using numbers or percentages over time.
- The process of creating an area plot involves importing Matplotlib and calling the plot function on the dataframe with ‘kind’ parameter assigned as ‘area.’
- Area plots provide a visually appealing and intuitive way to showcase the relationship and proportion of multiple variables in a single chart.

# Histograms

---

© IBM Corporation. All rights reserved.

# What you will learn

---



Define a histogram  
with the help of an  
illustration



Explore the process  
of creating a  
histogram using  
Matplotlib

# Histogram

---

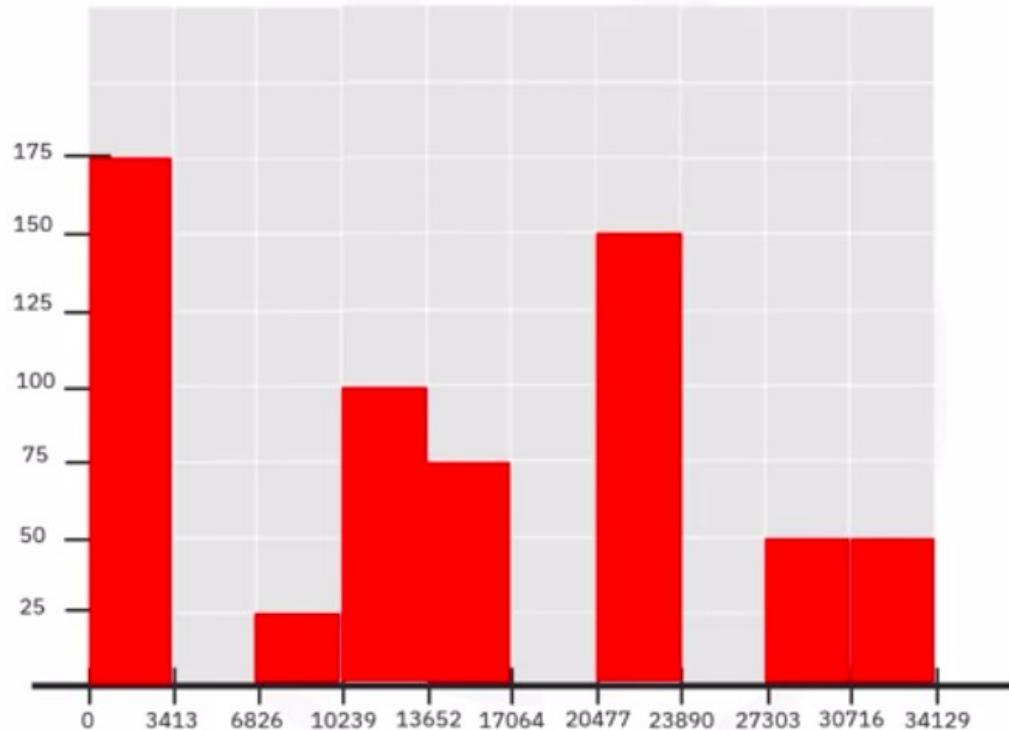
Represents the frequency distribution of a dataset

- Partitions numeric data into bins
- Assigns each data point in the dataset
- Counts the number of data points



# Histogram

---



# Dataset: Processed

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Country																					
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	6	
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

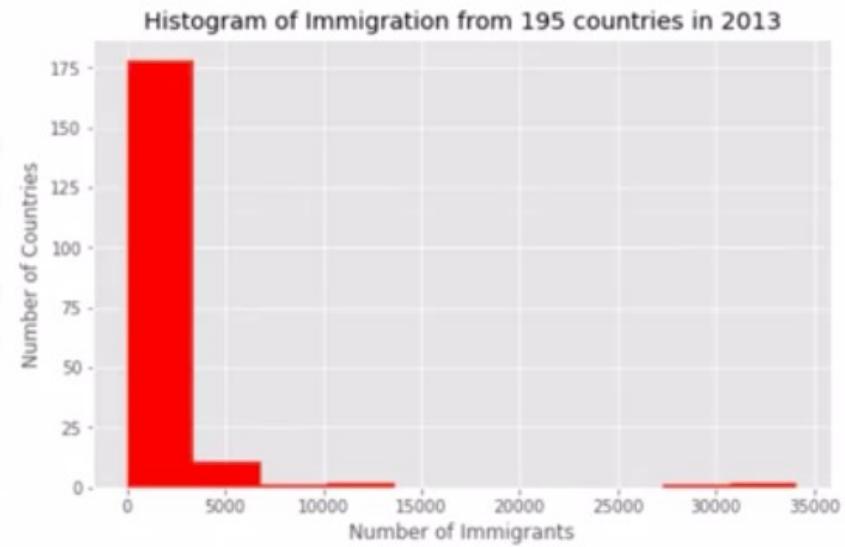
df\_canada

# Generating a histogram

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
<b>Country</b>																					
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	6	
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

# Generating a histogram

```
import matplotlib as mpl  
import matplotlib.pyplot as plt  
  
df_canada['2013'].plot(kind='hist')  
  
plt.title('Histogram of Immigration from 195 countries in 2013')  
plt.ylabel('Number of Countries')  
plt.xlabel('Number of Immigrants')  
  
plt.show()
```



# Generating a histogram

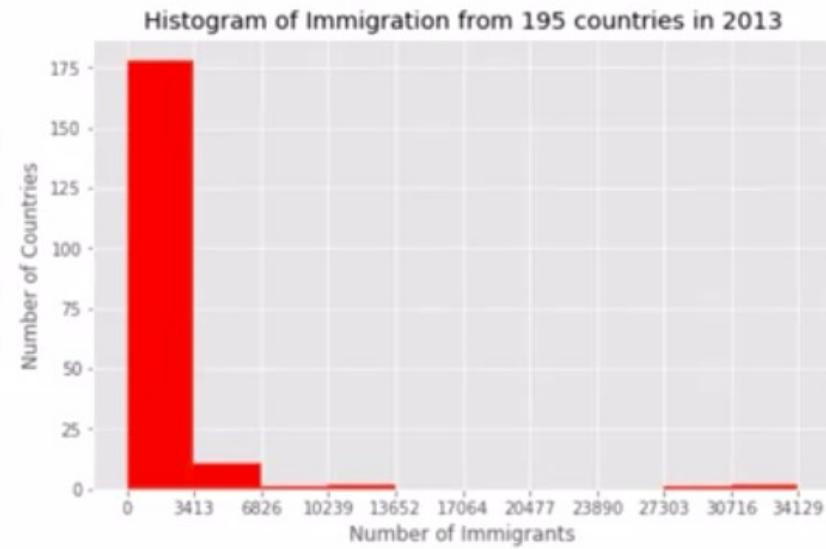
```
import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np

count, bin_edges = np.histogram(df_canada['2013'])

df_canada['2013'].plot(kind='hist', xticks = bin_edges)

plt.title('Histogram of Immigration from 195 countries in 2013')
plt.ylabel('Number of Countries')
plt.xlabel('Number of Immigrants')

plt.show()
```



# Recap

---

In this video, you learned that:

- A histogram is a way of representing the frequency distribution of a numeric dataset.
- To generate a histogram on Matplotlib, you import Matplotlib as mpl and its scripting interface as plt.
- You can call the plot function on the dataframe with kind parameter assigned as ‘hist.’
- You can use the Numpy library to create bins for the histogram representation.

# Bar Charts

---

© IBM Corporation. All rights reserved.

# What you will learn

---



Describe a bar chart with the help of an illustration

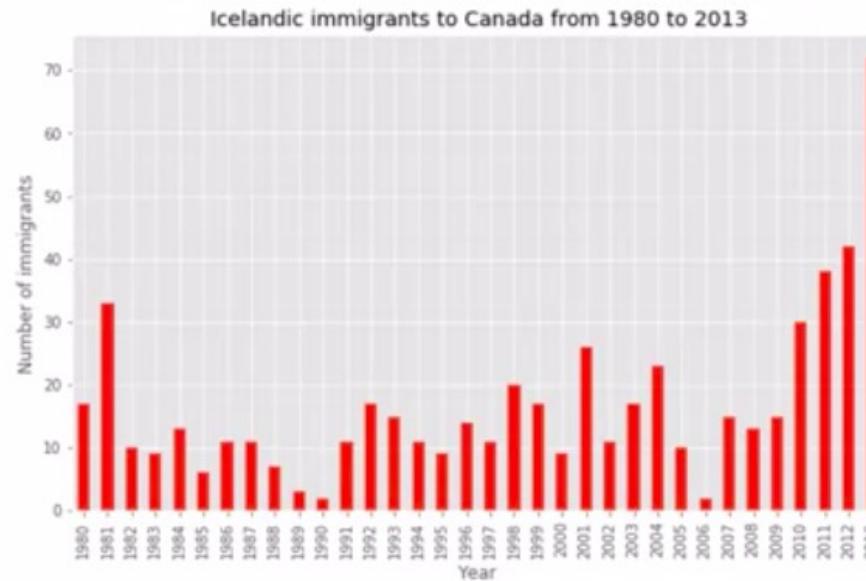


Explore the process of creating a bar chart using Matplotlib

# Bar chart

---

- Also known as a bar graph
- It compares the values of a variable



# Generating a bar chart

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
<b>Country</b>																					
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	0	6
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

df\_canada

# Generating a bar chart

```
import matplotlib as mpl
import matplotlib.pyplot as plt

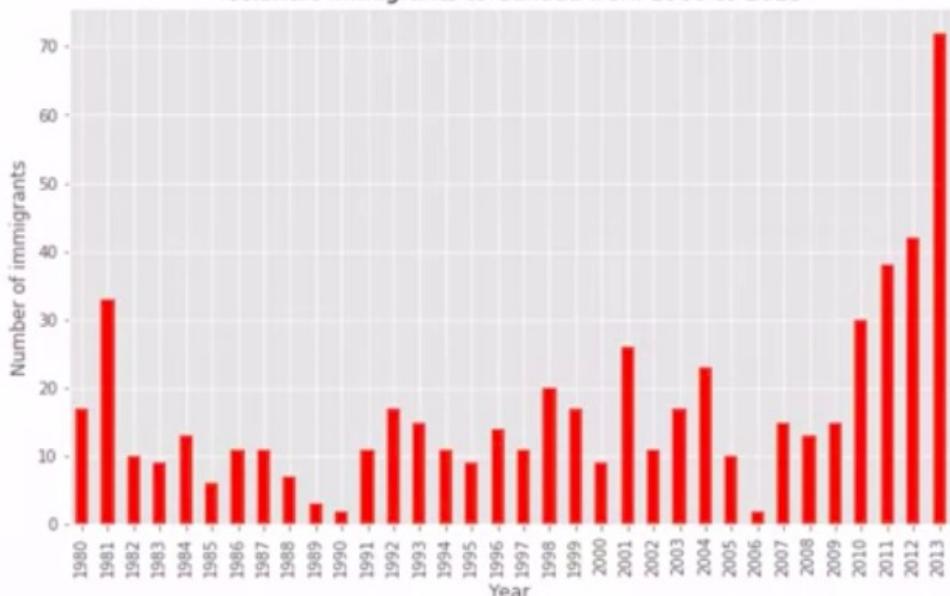
Years = list(map(str, range(1980, 2014)))
df_iceland = df_canada.loc[ 'Iceland',
years]

df_iceland.plot(kind='bar')

plt.title('Icelandic immigrants to
Canada from 1980 to 2013')
plt.xlabel('Year')
plt.ylabel('Number of immigrants')

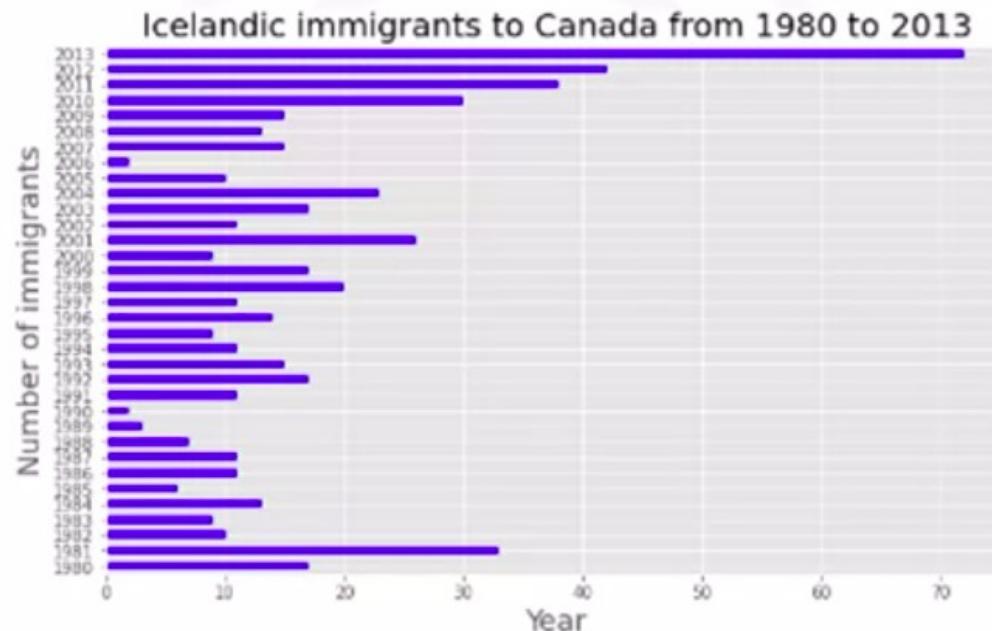
plt.show()
```

Icelandic immigrants to Canada from 1980 to 2013



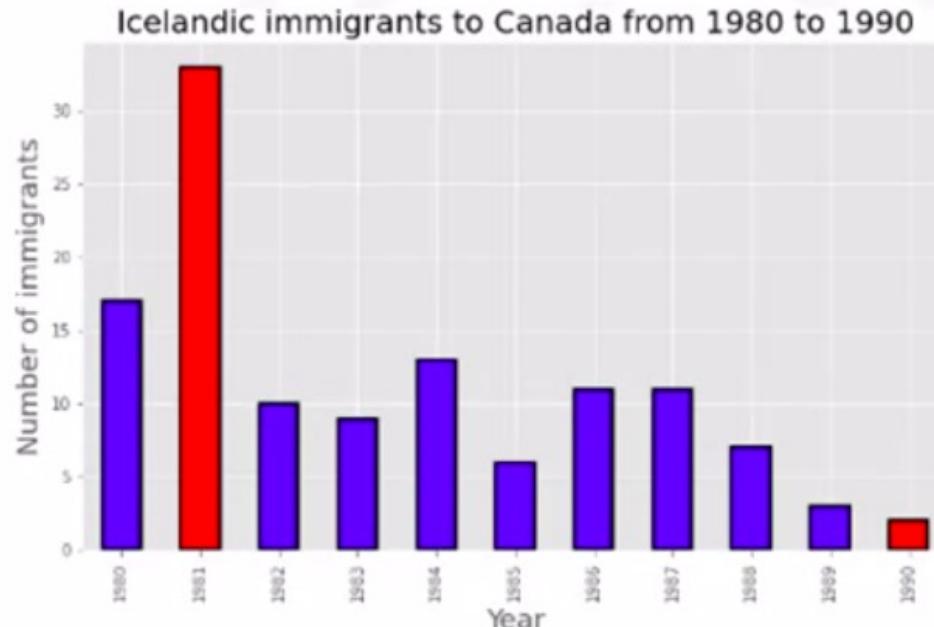
# Horizontal bar chart

```
df_iceland.plot(kind='barh', color = 'red')
```



# Highlighted bar and edge color

```
c=['blue','red',"blue",'blue','blue','blue','blue','blue','blue','blue','red']  
df_Iceland.plot(kind = 'bar', color = c, edgecolor = 'black')
```



# Recap

---

In this video, you learned that:

- A bar chart is a type of plot where the length of each bar is proportional to the value of the item that it represents.
- You can create a bar chart using Matplotlib representing the total immigration from Iceland to Canada.

A blue left-pointing arrow icon.

## Practice Quiz: Basic Visualization Tools

Practice Quiz • 10 min • 5 total points

## A green checkmark icon followed by the text "Congratulations! You passed!"

Grade received 100% To pass 80% or higher

[Go to next item](#)

1. **Select two:** An area plot depicts cumulated totals using \_\_\_\_\_ or \_\_\_\_\_ over time.

1 / 1 point

- Charts
- Data
- Percentages

## Correct

Correct! An area plot depicts cumulated totals using percentages over time.

- Numbers

## Correct

Correct! An area plot depicts cumulated totals using numbers over time.

A blue left-pointing arrow icon.

## Practice Quiz: Basic Visualization Tools

Practice Quiz • 10 min • 5 total points

2. An area plot, also known as an area chart or graph, displays the \_\_\_\_\_ of multiple variables.

1 / 1 point

- Magnitude and area
- Proportion and perimeter
- Proportion and area
- Magnitude and proportion

A green checkmark icon inside a circle.

Correct

Correct! An area plot, also known as an area chart or graph, displays the magnitude and proportion of multiple variables over a continuous axis, typically representing time or another ordered dimension.

3. A histogram is a way of representing the frequency distribution of a \_\_\_\_\_.

1 / 1 point

- Numeric dataset
- Alphabetical dataset
- Demographic dataset

A blue back arrow icon.

## Practice Quiz: Basic Visualization Tools

Practice Quiz • 10 min • 5 total points

Correct! An area plot, also known as an area chart or graph, displays the magnitude and proportion of multiple variables over a continuous axis, typically representing time or another ordered dimension.

3. A histogram is a way of representing the frequency distribution of a \_\_\_\_\_.

1 / 1 point

- Numeric dataset
- Alphabetical dataset
- Demographic dataset
- Statistical dataset

A green circular icon with a white checkmark inside.

Correct

Correct! The way a histogram works is that it partitions the spread of the numeric data into bins, assigns each data point in the dataset to a bin, and then counts the number of data points assigned to each bin.

4. In a Histogram, the \_\_\_\_\_ axis is the frequency or the number of data points in each bin.

1 / 1 point

- Perpendicular
- Vertical

A blue back arrow icon.

## Practice Quiz: Basic Visualization Tools

Practice Quiz • 10 min • 5 total points

Correct! The way a histogram works is that it partitions the spread of the numeric data into bins, assigns each data point in the dataset to a bin, and then counts the number of data points assigned to each bin.

4. In a Histogram, the \_\_\_\_\_ axis is the frequency or the number of data points in each bin.

1 / 1 point

- Perpendicular
- Vertical
- Parallel
- Horizontal



Correct! The vertical axis is the frequency or the number of data points in each bin.

5. The following code will create a horizontal bar chart of the data in the panda's data frame **question**.

1 / 1 point

```
1 question.plot(type='bar', rot=90)
```

A blue left-pointing arrow icon.

## Practice Quiz: Basic Visualization Tools

Practice Quiz • 10 min • 5 total points



Correct! The vertical axis is the frequency or the number of data points in each bin.

5. The following code will create a horizontal bar chart of the data in the panda's data frame **question**.

1 / 1 point

```
1 question.plot(type='bar', rot=90)
```

 False True

Correct! The parameters used are wrong. For example, type is used instead of kind.

# Pie Charts

---

© IBM Corporation. All rights reserved.

# What you will learn

---



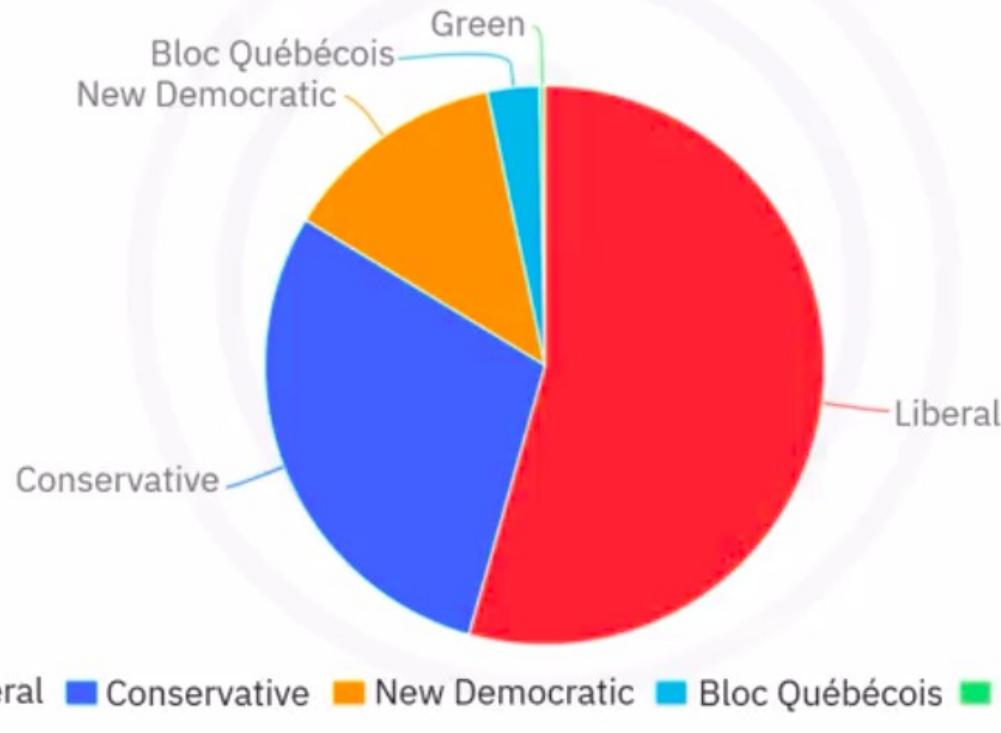
Describe a pie chart  
with the help of an  
example



Explore the process  
of creating a pie chart  
using Matplotlib

# Pie chart

A circular statistical graphic, divided into segments



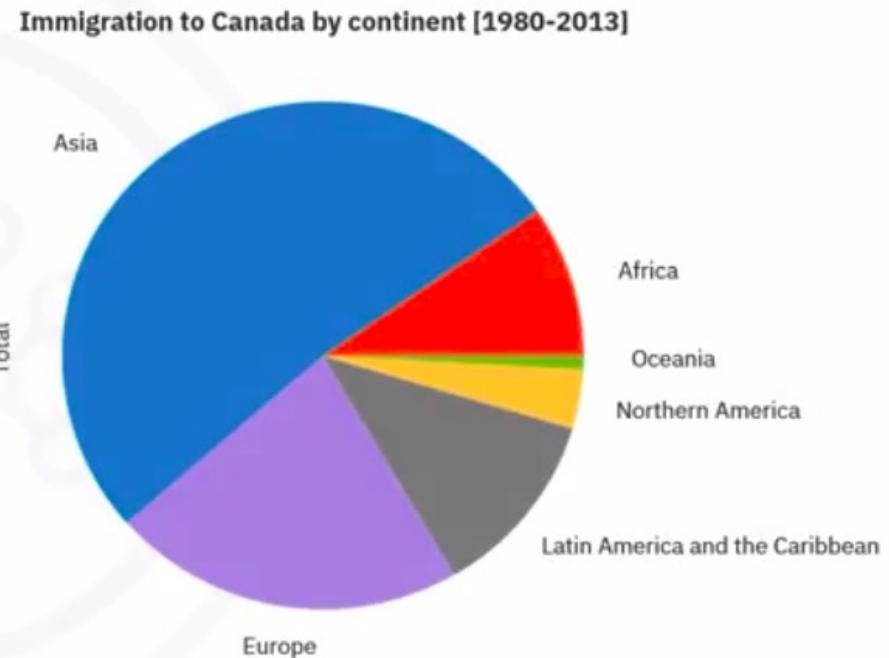
# Pie chart

```
df_continents = df_canada.groupby('Continent', axis = 0).sum()
```

	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Continent																					
Africa	3951	4363	3819	2671	2639	2650	3782	7494	7552	9894	...	27523	29188	28284	29890	34534	40892	35441	38083	38543	618948
Asia	31025	34314	30214	24696	27274	23850	28739	43203	47454	60256	...	159253	149054	133459	139894	141434	163845	146894	152218	155075	3317794
Europe	39760	44802	42720	24638	22287	20844	24370	46698	54726	60893	...	35955	33053	33495	34692	35078	33425	26778	29177	28691	1410947
Latin America and the Caribbean	13081	15215	16769	15427	13678	15171	21179	28471	21924	25060	...	24747	24676	26011	26547	26867	28818	27856	27173	24950	765148
Northern America	9378	10030	9074	7100	6661	6543	7074	7705	6469	6790	...	8394	9613	9463	10190	8995	8142	7677	7892	8503	241142
Oceania	1942	1839	1675	1018	878	920	904	1200	1181	1539	...	1585	1473	1693	1834	1860	1834	1548	1679	1775	55174

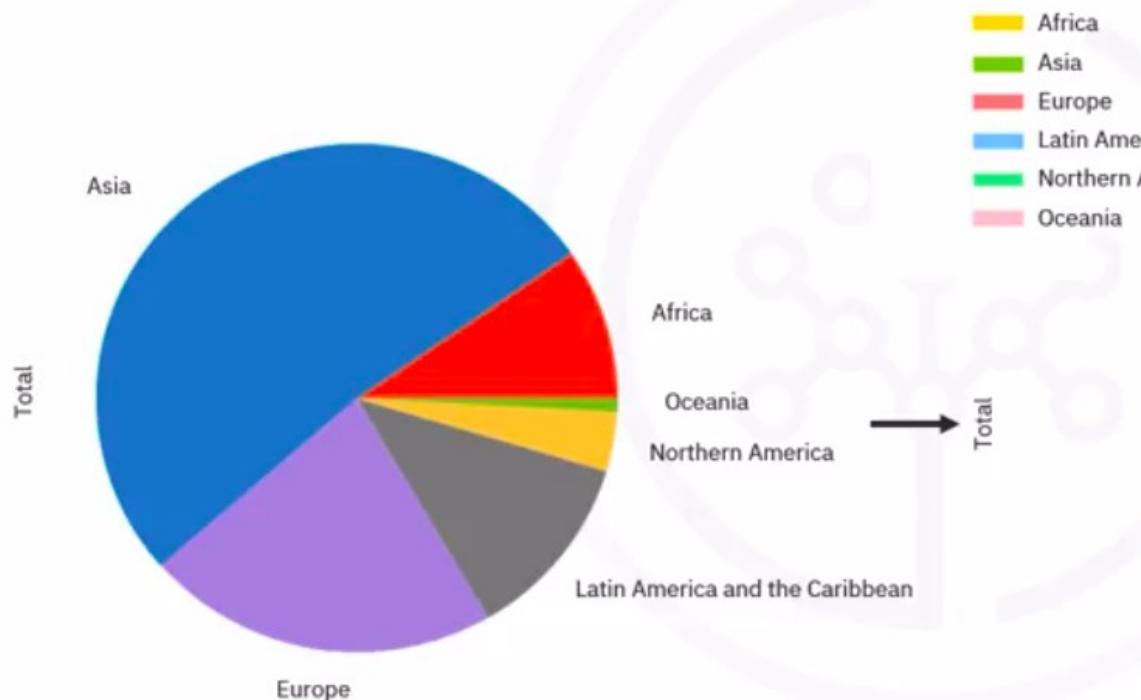
# Pie chart

```
import matplotlib as mpl  
import matplotlib.pyplot as plt  
  
df_continents['Total'].plot(kind='pie')  
  
plt.title('Immigration to Canada by  
Continent [1980-2013]')  
  
plt.show()
```

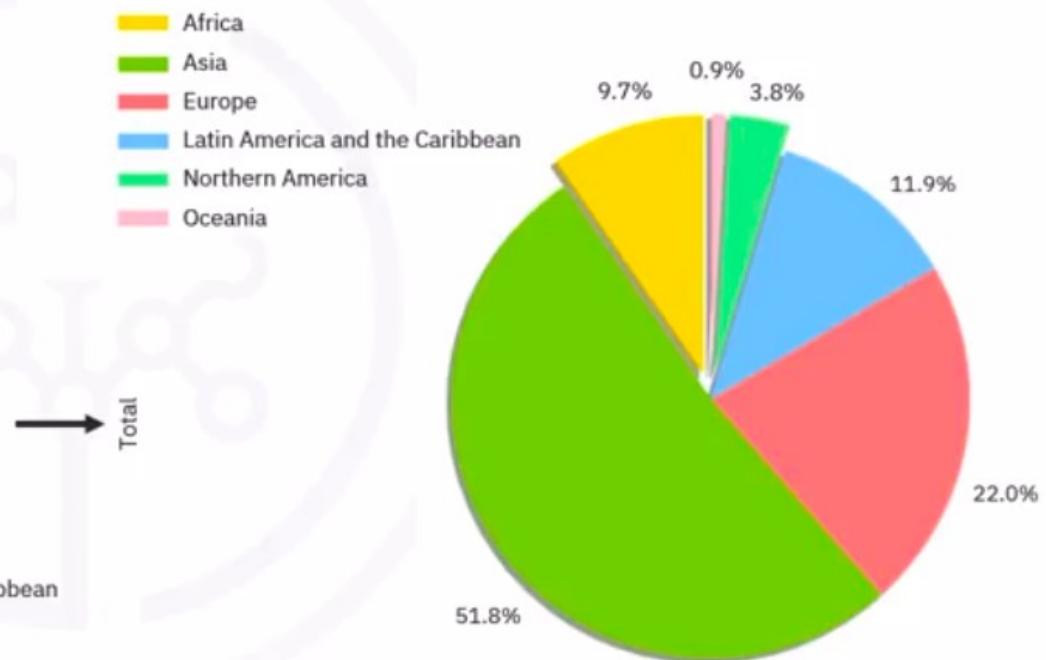


# Pie chart: Explode property

Immigration to Canada by continent [1980-2013]



Immigration to Canada by continent [1980-2013]



# Pie chart: Flaws

---

Link:

<https://www.surveygizmo.com/survey-blog/pie-chart-or-bar-graph/>

# Recap

---

In this video, you learned that:

- A pie chart is a circular statistical graphic, divided into segments, to illustrate numerical proportion.
- The process of creating a pie chart involves importing Matplotlib to represent a large set of data over a period of time.

# Box Plots

---

© IBM Corporation. All rights reserved.

# What you will learn

---

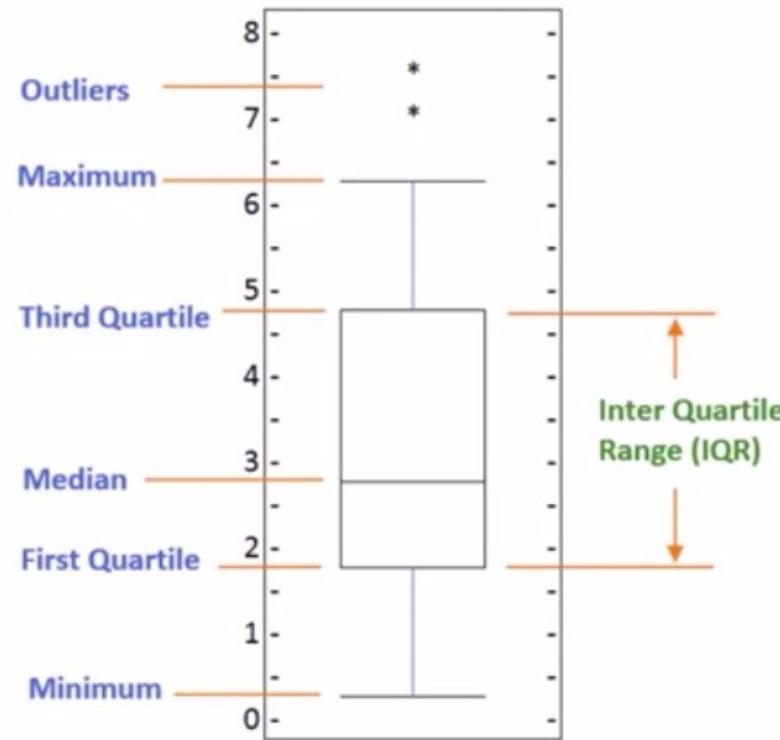


Describe a box plot  
with the help of an  
illustration



Explain how to create  
a box plot using  
Matplotlib

# Box plot



# Generating a box plot

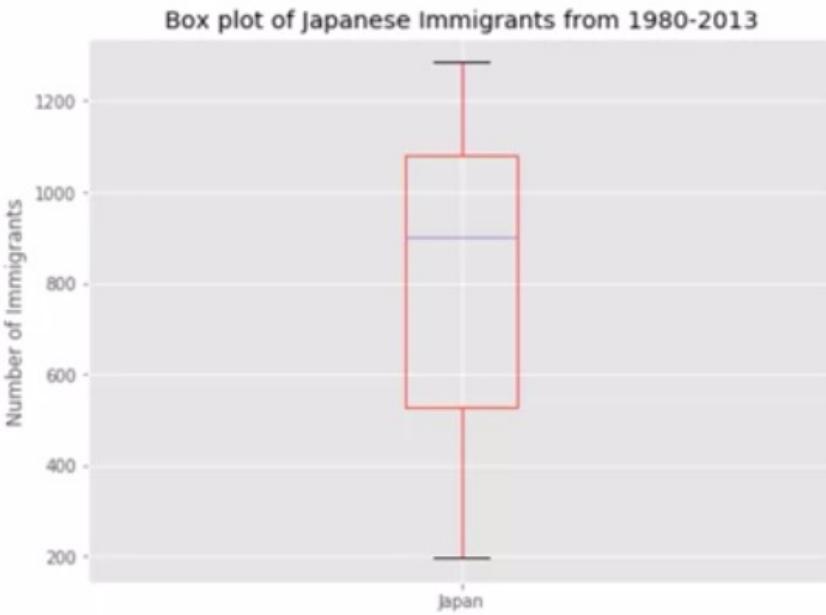
---

- Process the data frame df\_canada

```
import matplotlib as mpl  
import matplotlib.pyplot as plt
```

# Generating a box plot

```
df_japan = df_canada.loc[['Japan'], years].transpose()  
  
df_japan.plot(kind='box')  
  
plt.title('Box plot of Japanese Immigrants from 1980-2013')  
plt.ylabel("Number of Immigrants")  
  
plt.show()
```



# Recap

---

In this video, you learned that:

- A box plot is a way of statistically representing given data distribution through five main dimensions.
- The five main dimensions are minimum, first quartile, median, third quartile, and maximum.
- You can create a box plot using Matplotlib.

# Scatter Plots

---

© IBM Corporation. All rights reserved.

# What you will learn

---



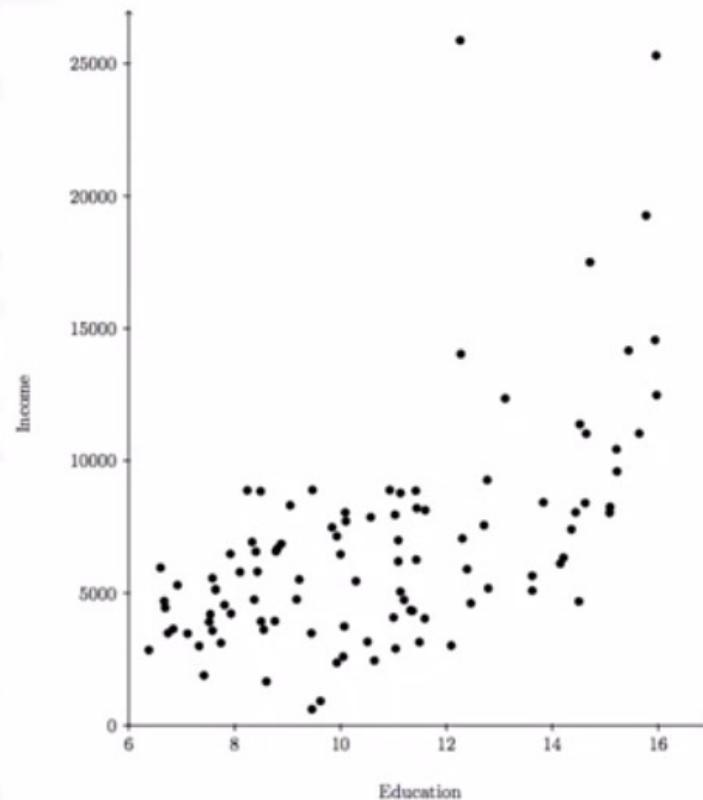
Describe what is a scatter plot with the help of an example



Explore the scatter plot creation process using Matplotlib

# Scatter plot

- Displays values pertaining to two variables
- Determines the correlation between the two variables



# Dataset: Processed

	Continent	Region	DevName	1980	1981	1982	1983	1984	1985	1986	...	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
<b>Country</b>																					
Afghanistan	Asia	Southern Asia	Developing regions	16	39	39	47	71	340	496	...	3436	3009	2652	2111	1746	1758	2203	2635	2004	58639
Albania	Europe	Southern Europe	Developed regions	1	0	0	0	0	0	1	...	1223	856	702	560	716	561	539	620	603	15699
Algeria	Africa	Northern Africa	Developing regions	80	67	71	69	63	44	69	...	3626	4807	3623	4005	5393	4752	4325	3774	4331	69439
American Samoa	Oceania	Polynesia	Developing regions	0	1	0	0	0	0	0	...	0	1	0	0	0	0	0	0	6	
Andorra	Europe	Southern Europe	Developed regions	0	0	0	0	0	0	2	...	0	1	1	0	0	0	0	1	1	15

df\_canada

# Generating a scatter plot

---

year	total
1980	99137
1981	110563
1982	104271
1983	75550
1984	73417
.	.
.	.

df\_total

# Generating a scatter plot

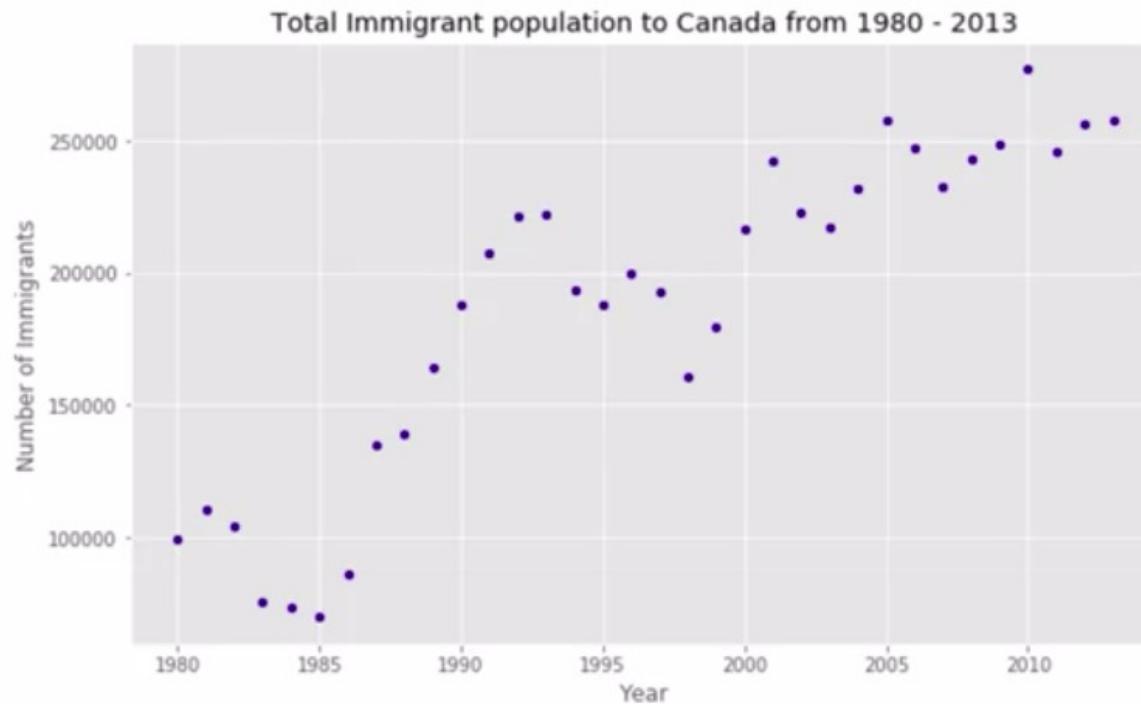
```
import matplotlib as mpl  
import matplotlib.pyplot as plt
```

```
df_total.plot(  
    kind='scatter',  
    x='year',  
    y='total',  
)  
  
plt.title('Total Immigrant population to Canada from 1980 - 2013')  
plt.xlabel ('Year')  
plt.ylabel('Number of Immigrants')  
  
plt.show()
```

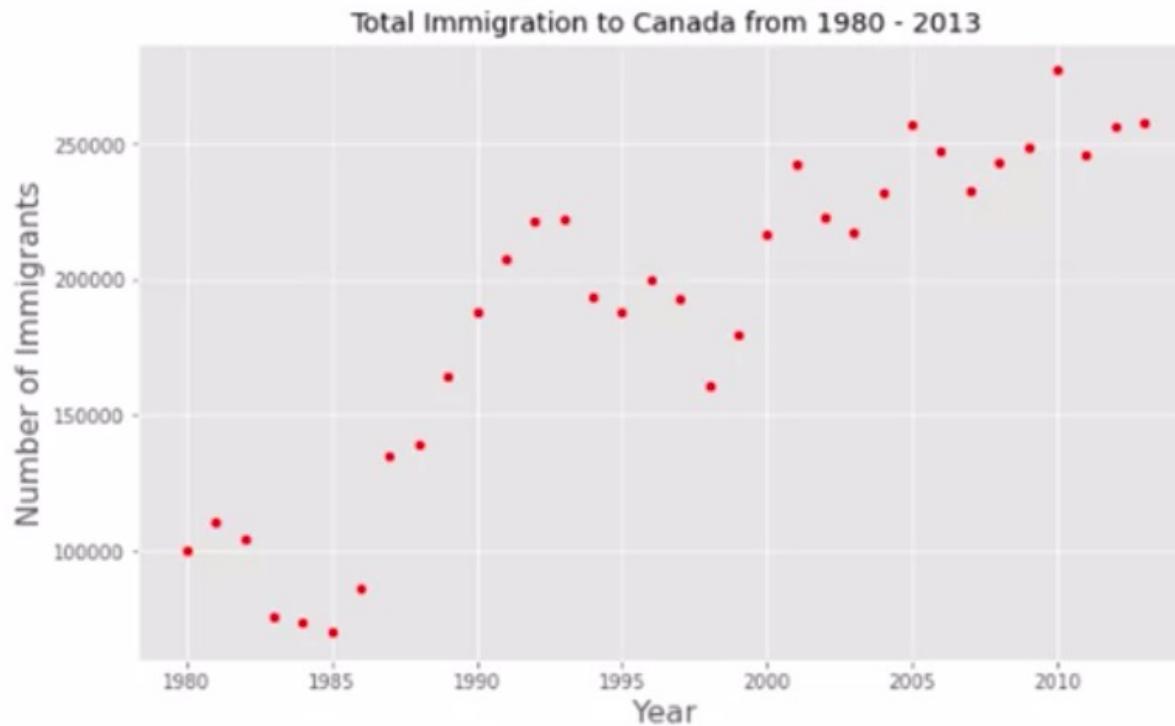
df\_total

year	total
1980	99137
1981	110563
1982	104271
1983	75550
1984	73417
.	.
.	.

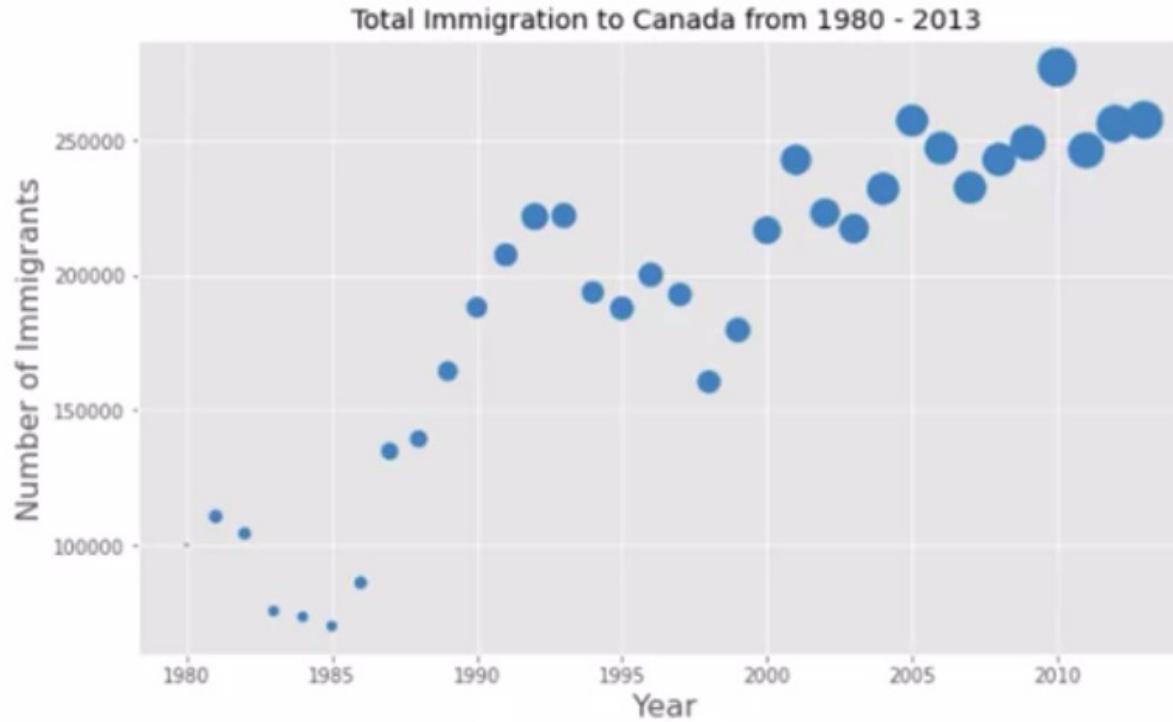
# Generating a scatter plot



# Generating a scatter plot



# Generating a scatter plot



# Recap

---

In this video, you learned that:

- A scatter plot displays values pertaining to typically two variables against each other.
- The process of creating a scatter plot using Matplotlib to visualize a large set of data.

# Plotting Directly with Matplotlib

---

© IBM Corporation. All rights reserved.

# What you will learn

---



Explore various functions offered by Matplotlib for data visualization and plotting



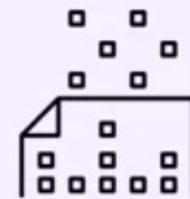
Differentiate between data storytelling and data visualization

# Importing libraries, figures, and axes

```
import matplotlib.pyplot as plt  
import numpy as np
```

```
import pandas as pd
```

```
# Create figure and axes  
fig, ax = plt.subplots()
```

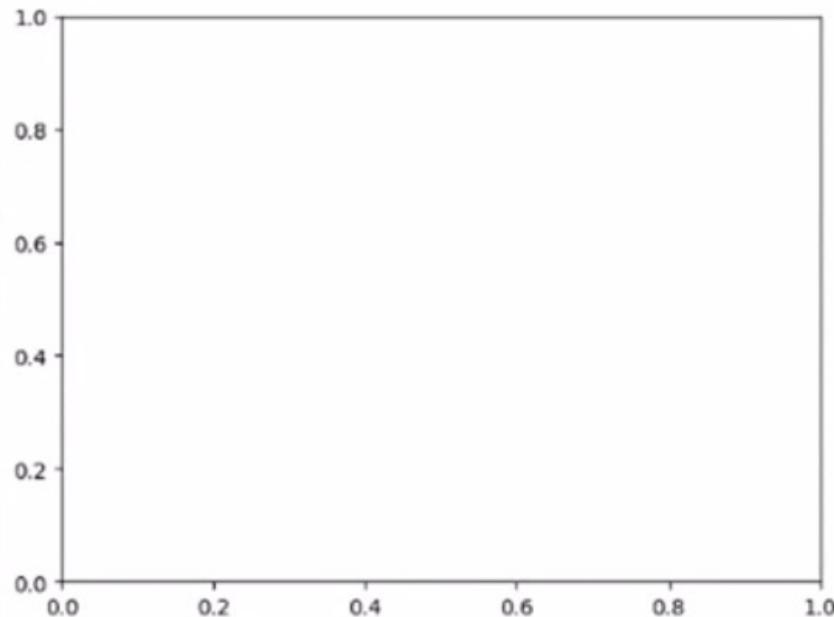


# Importing libraries, figures, and axes

```
import matplotlib.pyplot as plt  
import numpy as np
```

```
import pandas as pd
```

```
# Create figure and axes  
fig, ax = plt.subplots()
```

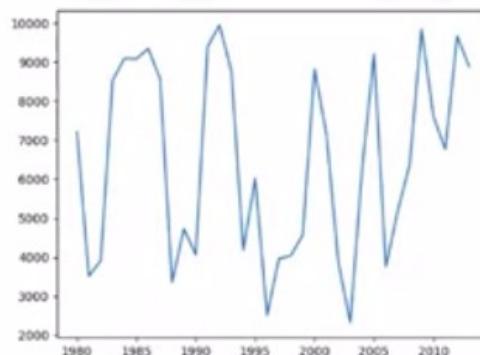


# Line plot and scatter plot

```
# Synthetic data  
years = np.arange(1980, 2014)  
immigrants = np.random.randint(2000, 10000, size = (34,))
```

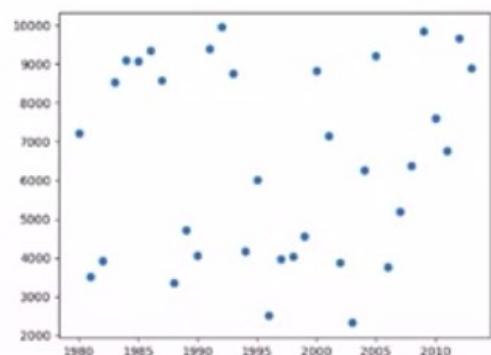
- Line Plot

```
#Create figure and axes  
fig, ax = plt.subplots()  
  
# Plot the Line  
ax.scatter(years, immigrants)  
  
# Display the plot  
plt.show()
```



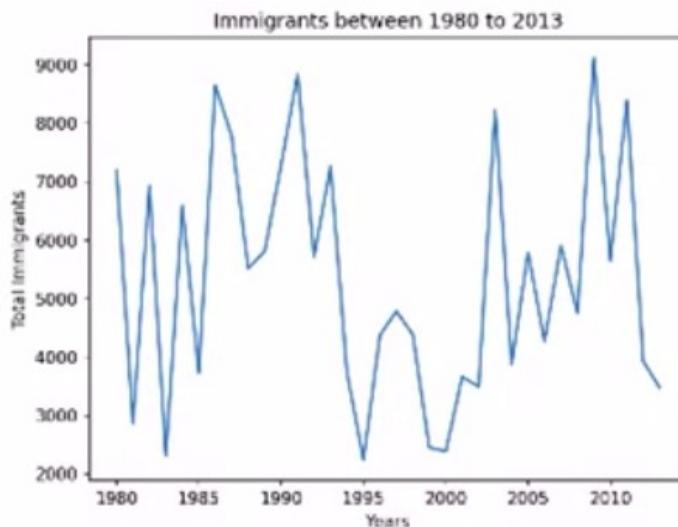
- Scatter Plot

```
#Create figure and axes  
fig, ax = plt.subplots()  
  
# Plot the Line  
ax.plot(years, immigrants)  
  
# Display the plot  
plt.show()
```

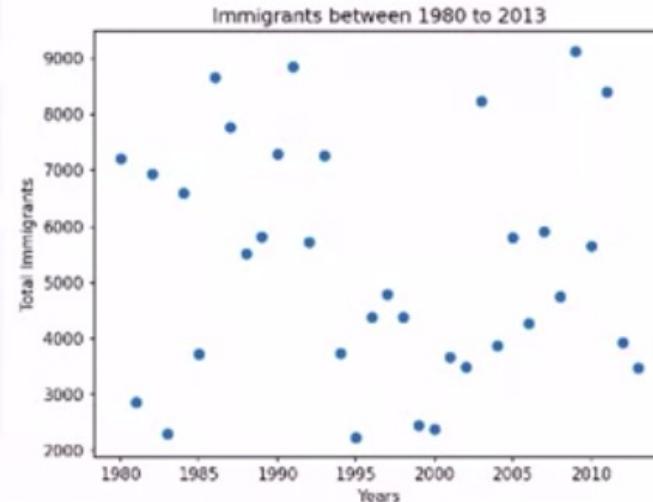


# Customizing attributes

```
#add title  
plt.title('Immigrants between 1980 to 2013')  
#add Labels  
plt.xlabel('Years')  
plt.ylabel('Total Immigrants')
```

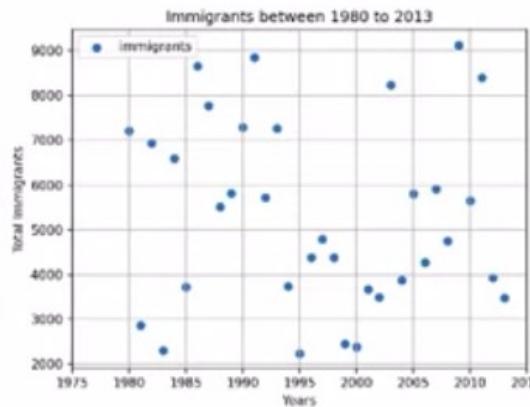
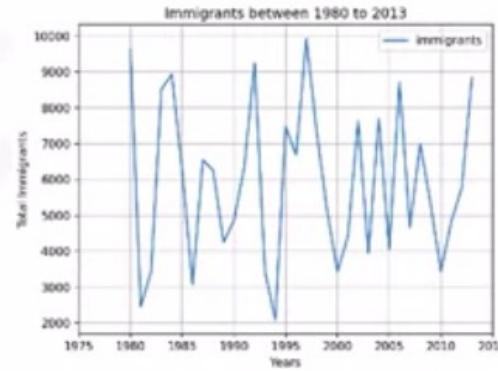


```
ax.set_title('Immigrants between 1980 to 2013')  
ax.set_xlabel('Years')  
ax.set_ylabel('Total Immigrants')
```



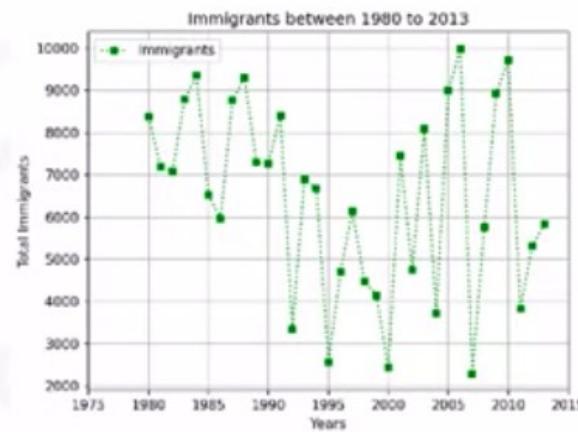
# Customizing attributes

```
#limits on x-axis  
plt.xlim(1975, 2015)  
#or ax.set_xlim()  
  
#Enabling Grid  
plt.grid(True)  
#or ax.grid()  
  
#Legend  
plt.legend(["Immigrants"]) #or ax.legend()
```



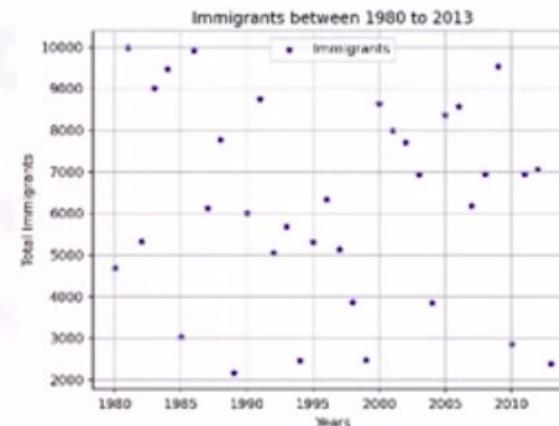
# Marker styles, colors, and sizes

```
# Customizing the appearance of Plot  
ax.plot(years, immigrants,  
        marker='s',  
        markersize=5,  
        color='green',  
        linestyle="dotted")
```



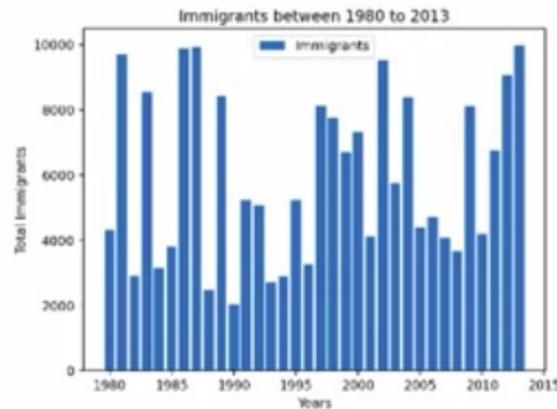
```
# Customizing Scatter Plot  
ax.scatter(years, immigrants,  
          marker='*',  
          s = 20,  
          color='darkblue')
```

```
#Legend at upper center of the figure  
ax.legend(["Immigrants"], loc='upper center')
```



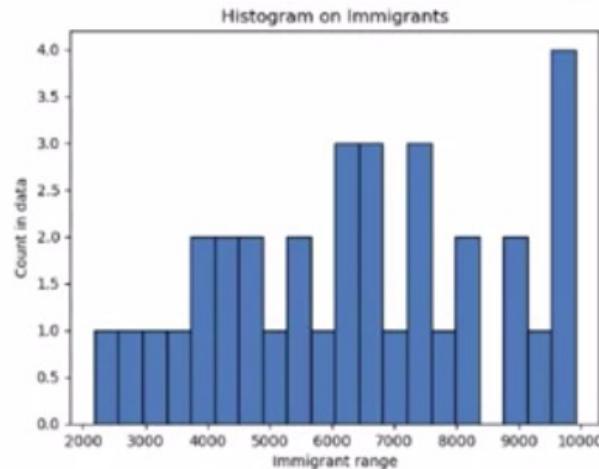
# Bar plot, histogram, and Pie

```
# Plot the bar  
ax.bar(years, immigrants)
```



# Bar plot, histogram, and Pie

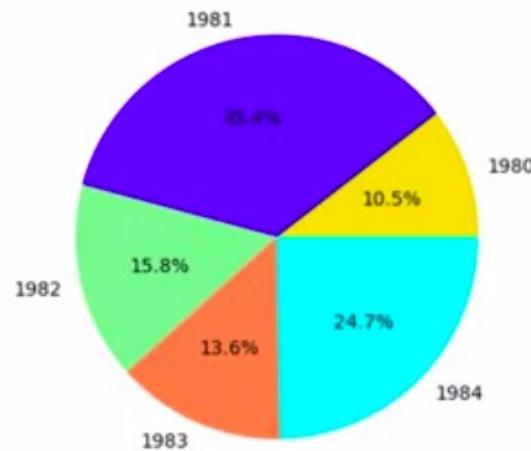
```
#histogram on immigrants  
ax.hist(immigrants, bins=20, edgecolor='black',color='steelblue')
```



# Bar plot, histogram, and Pie

```
#Pie on immigrants  
ax.pie(immigrants[0:5], labels=years[0:5],  
       colors = ['gold','blue','lightgreen','coral','cyan'],  
       autopct='%1.1f%%')
```

Distribution of Immigrants Over Years

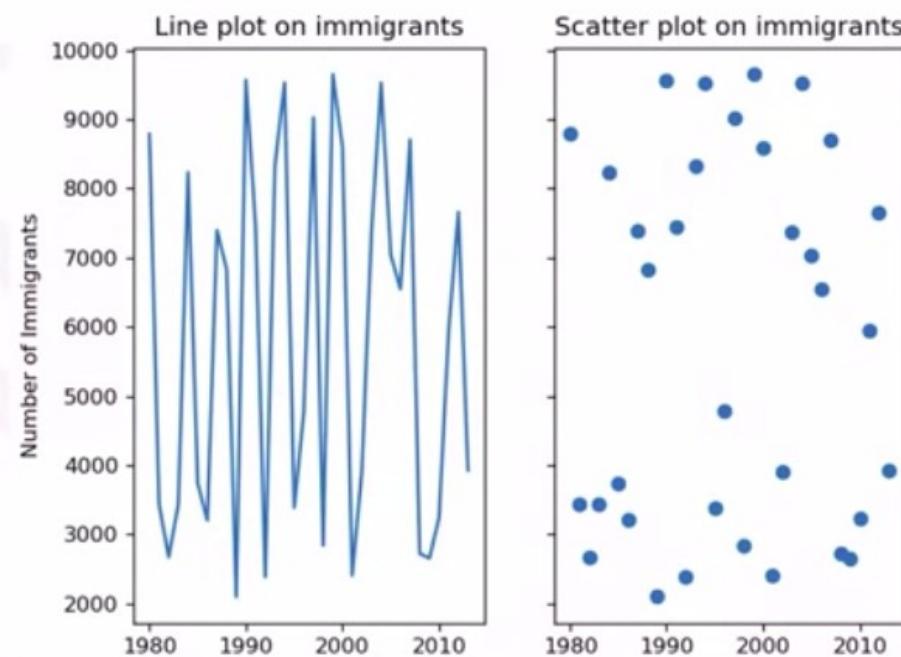


# Multiple plots and sub-plotting

```
# Create a figure with two axes in a row
fig, axs = plt.subplots(1, 2, sharey=True)

#Plotting in first axes - the left one
axs[0].plot(years, immigrants)
axs[0].set_title("Line plot on immigrants")

#Plotting in second axes - the right one
axs[1].scatter(years, immigrants)
axs[1].set_title("Scatter plot on immigrants")
```



# Multiple plots and sub-plotting

- Three Arguments
  - Number of rows
  - Columns
  - Index of the subplot

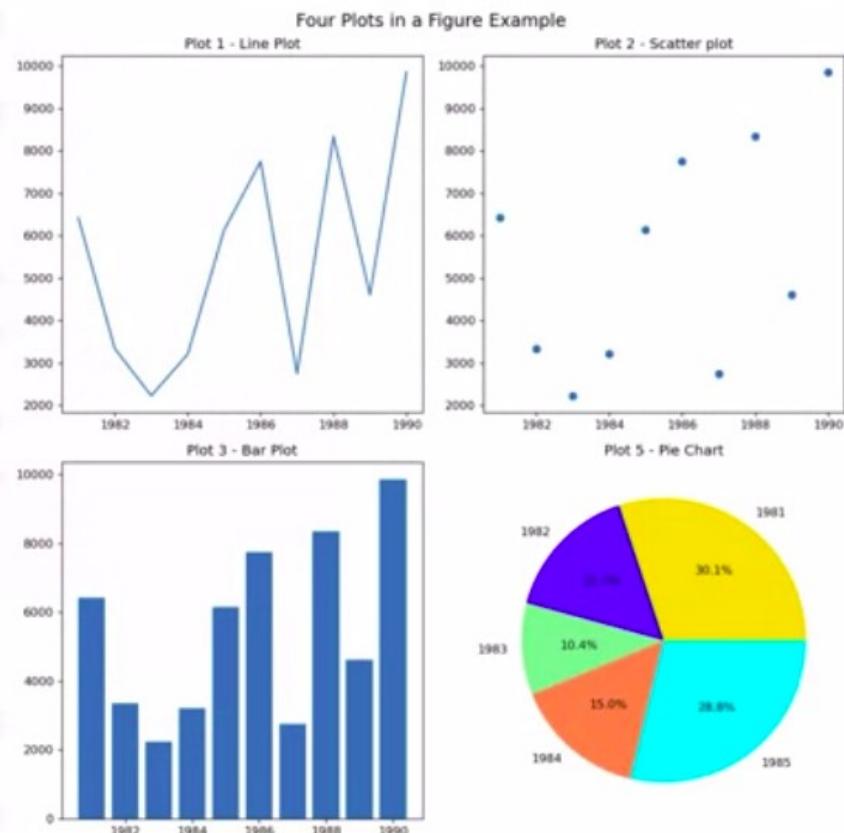
```
# Add the first subplot (top-left)
axs1 = fig.add_subplot(2, 2, 1)
#Plotting in first axes - the left one
axs1.plot(total_immigrants)
axs1.set_title("Plot 1 - Line plot")
```

# Multiple plots and sub-plotting

```
# Add the second subplot (top-right)
ax2 = fig.add_subplot(2, 2, 2)
ax2.scatter(years, immigrants)
ax2.set_title('Plot 2 - Scatter plot')

# Add the third subplot (bottom-left)
ax3 = fig.add_subplot(2, 2, 3)
ax3.bar(years, immigrants)
ax3.set_title('Plot 3 - Bar Plot')

# Add the fourth subplot (bottom-right)
ax4 = fig.add_subplot(2, 2, 4)
ax4.pie(immigrants[0:5], labels=years[0:5],
        colors = ['gold','blue','lightgreen','coral','cyan'],
        autopct='%.1f%%')
ax4.set_aspect('equal')
ax4.set_title('Plot 5 - Pie Chart')
```



# Data storytelling and data visualization

---

- Data storytelling
  - Is the art of storytelling
  - Creates a narrative around the data
- Data visualization
  - Is an important aspect of data storytelling
  - Creates engaging visuals

## Article link:

<https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/?sh=5aca60cd52ad>

# Recap

---

In this video, you learned that:

- Matplotlib is a versatile plotting library that offers a flexible interface for creating various types of plots.
- Matplotlib's pyplot module offers a convenient way to create and customize plots quickly.
- Data Storytelling is the ‘art of storytelling’ that involves creating a narrative around the data.

# Recap

---

In this video, you learned that:

- Matplotlib is a versatile plotting library that offers a flexible interface for creating various types of plots.
- Matplotlib's pyplot module offers a convenient way to create and customize plots quickly.
- Data Storytelling is the ‘art of storytelling’ that involves creating a narrative around the data.
- Data visualization is an important aspect of data storytelling and involves creating engaging visuals.

[Back](#) Practice Quiz: Specialized Visualization Tools  
Practice Quiz • 10 min • 5 total points

## Congratulations! You passed!

Grade received 100% To pass 80% or higher

[Go to next item](#)

1. What is a scatter plot?

1 / 1 point

- A scatter plot is a type of plot that displays values pertaining to typically two variables against each other
- A scatter plot is a circular statistical graphic divided into segments
- A scatter plot displays the magnitude and proportion of multiple variables over a continuous axis
- A scatter plot represents the frequency distribution of a numeric dataset

Correct

Correct! A scatter plot is a type of plot that displays values pertaining to typically two variables against each other. Usually, it is a dependent variable plotted against an independent variable.

2. True or False. Matplotlib is a general-purpose comprehensive plotting library that provides a flexible interface for creating a wide range of plots.

1 / 1 point

[Back](#)

## Practice Quiz: Specialized Visualization Tools

Practice Quiz • 10 min • 5 total points

2. True or False. Matplotlib is a general-purpose comprehensive plotting library that provides a flexible interface for creating a wide range of plots.

1 / 1 point

- True  
 False

 **Correct**

Correct! Matplotlib is a general-purpose comprehensive plotting library that provides a flexible interface for creating a wide range of plots. Its pyplot module offers a convenient way to create and customize plots quickly.

3. What is a pie chart?

1 / 1 point

- A pie chart is a type of plot in which the length of each bar is proportional to the value of the item that it represents  
 A pie chart is a graphical representation that showcases the relative size and proportion of various variables along a continuous axis  
 A pie chart is a circular statistical graphic divided into segments to illustrate numerical proportions  
 A pie chart visually depicts the distribution of a numeric dataset by showcasing the frequency of each category

 **Correct**

Correct! A pie chart is a circular statistical graphic divided into segments to illustrate numerical proportions. The explode property in a pie chart enables you to offset slices from the center, highlighting specific sections.

[Back](#) Practice Quiz: Specialized Visualization Tools  
Practice Quiz • 10 min • 5 total points

offset slices from the center, highlighting specific sections.

4. A box plot is a way of statistically representing the distribution of given data through how many primary dimensions?

1 / 1 point

- 8
- 5
- 3
- 1



Correct

Correct! A box plot is a way of statistically representing the distribution of given data through five main dimensions. These include Minimum, First quartile, Median, Third quartile, and Maximum.

5. What is the first step when plotting with Matplotlib?

1 / 1 point

- Call the plot function
- Import Pandas

Back

## Practice Quiz: Specialized Visualization Tools

Practice Quiz • 10 min • 5 total points

3

1

Correct

Correct! A box plot is a way of statistically representing the distribution of given data through five main dimensions. These include Minimum, First quartile, Median, Third quartile, and Maximum.

5. What is the first step when plotting with Matplotlib?

1 / 1 point

Call the plot function

Import Pandas

Call the subplot function

Import matplotlib.pyplot as plt

Correct

Correct! The first step is to import the library. You import matplotlib.pyplot as plt.



Search in course

Search



Tushar Raha

Data Visualization with Python > Week 2 > Summary: Basic and Specialized Visualization Tools

< Previous Next >

✓ Video: Plotting Directly with Matplotlib

8 min

✓ Ungraded App Item: Hands-on Lab: Plotting Directly with Matplotlib

40 min

✓ Practice Quiz: Practice Quiz: Specialized Visualization Tools

5 questions

Reading: Summary: Basic and Specialized Visualization Tools

5 min

Ungraded Plugin: Cheat Sheet: Plotting with Matplotlib using Pandas

15 min

Quiz: Graded Quiz: Basic and Specialized Visualization Tools

10 questions

Congratulations! You have completed this module. At this point in the course, you know:

- A pie chart is a circular statistical graphic, divided into segments, to illustrate numerical proportion.
- The process of creating a pie chart involves importing Matplotlib to represent a large set of data over a period of time.
- A box plot is a way of statistically representing given data distribution through five main dimensions.
- The five main dimensions are minimum, first quartile, median, third quartile, and maximum.
- You can create a box plot using Matplotlib.
- A scatter plot displays values pertaining to typically two variables against each other.
- The process of creating a scatter plot involves importing Matplotlib to visualize a large set of data.
- Matplotlib is a versatile plotting library that offers a flexible interface for creating various types of plots.
- Matplotlib's Pyplot module offers a convenient way to create and customize plots quickly.
- Data Storytelling is the 'art of storytelling' that involves creating a narrative around the data.
- Data visualization is an important aspect of data storytelling and involves creating engaging visuals.



[Back](#)

## Graded Quiz: Basic and Specialized Visualization Tools

Due Oct 8, 11:59 PM IST

Graded Quiz • 30 min

## Congratulations! You passed!

[Go to next item](#)

Grade received 100% Latest Submission Grade 100% To pass 70% or higher

1. What does a scatter plot display?

1 / 1 point

- Data
- Values pertaining to typically two variables against each other.
- Graphs
- Numbers



Correct

Correct! A scatter plot displays values pertaining to typically two variables against each other. Usually, it is a dependent variable that is plotted against an independent variable to determine if any correlation between the two variables exists.

2. When plotting directly with Matplotlib, what module offers a convenient way to create and customize plots quickly?

1 / 1 point

[Back](#)

## Graded Quiz: Basic and Specialized Visualization Tools

Due Oct 8, 11:59 PM IST

Graded Quiz • 30 min



Correct

Correct! A scatter plot displays values pertaining to typically two variables against each other. Usually, it is a dependent variable that is plotted against an independent variable to determine if any correlation between the two variables exists.

2. When plotting directly with Matplotlib, what module offers a convenient way to create and customize plots quickly?

1 / 1 point

- Folium
- Numpy
- Pyplot
- Plotly



Correct

Correct! Matplotlib is a general-purpose comprehensive plotting library. Its pyplot module offers a convenient way to create and customize plots quickly.

3. A pie chart is a \_\_\_\_\_ statistical graphic, divided into segments, to illustrate numerical proportions.

1 / 1 point

- Circular

[Back](#)

## Graded Quiz: Basic and Specialized Visualization Tools

Graded Quiz • 30 min

Due Oct 8, 11:59 PM IST

 Plotly **Correct**

Correct! Matplotlib is a general-purpose comprehensive plotting library. Its pyplot module offers a convenient way to create and customize plots quickly.

3. A pie chart is a \_\_\_\_\_ statistical graphic, divided into segments, to illustrate numerical proportions.

1 / 1 point

- Circular
- Bar chart
- Line plot
- Folium

**Correct**

Correct! A pie chart is a circular statistical graphic, divided into segments, to illustrate numerical proportions.

4. How many primary dimensions does a box plot utilize to statistically represent the distribution of a given data?

1 / 1 point

[Back](#)

## Graded Quiz: Basic and Specialized Visualization Tools

Due Oct 8, 11:59 PM IST

Graded Quiz • 30 min

Correct! A pie chart is a circular statistical graphic, divided into segments, to illustrate numerical proportions.

4. How many primary dimensions does a box plot utilize to statistically represent the distribution of a given data?

1 / 1 point

- 1
- 4
- 3
- 5

Correct

Correct! A box plot is a way of statistically representing the distribution of given data through five primary dimensions. These include Minimum, First quartile, Median, Third quartile, and Maximum.

5. Area lots are similar to what other type of plot?

1 / 1 point

- Area
- Line

[Back](#)

## Graded Quiz: Basic and Specialized Visualization Tools

Due Oct 8, 11:59 PM IST

Graded Quiz • 30 min

Correct! A box plot is a way of statistically representing the distribution of given data through five primary dimensions. These include Minimum, First quartile, Median, Third quartile, and Maximum.

5. Area lots are similar to what other type of plot?

1 / 1 point

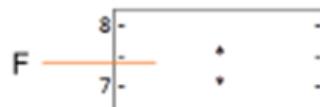
- Area
- Line
- Box
- Bar

Correct

Correct! An area plot, also known as an area chart or graph, displays the magnitude and proportion of multiple variables over a continuous axis, typically representing time or another ordered dimension.

6.

1 / 1 point



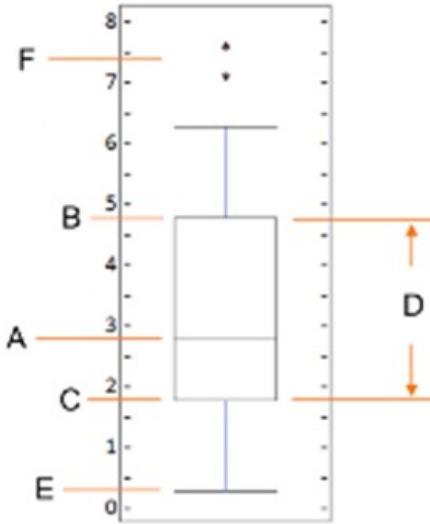
[Back](#)

## Graded Quiz: Basic and Specialized Visualization Tools

Due Oct 8, 11:59 PM IST

Graded Quiz • 30 min

6.



1 / 1 point

In the above chart, what do the letters in the box plot above represent?

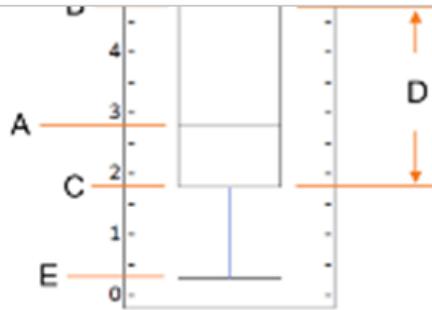
- A = Median, B = Third Quartile, C = First Quartile, D = Inter Quartile Range, E = Minimum, and F = Outliers
- A = Mean, B = Third Quartile, C = First Quartile, D = Inter Quartile Range, E = Minimum, and F = Maximum
- A = Mean, B = Upper Mean Quartile, C = Lower Mean Quartile, D = Inter Quartile Range, E = Minimum, and F = Outliers

Back

## Graded Quiz: Basic and Specialized Visualization Tools

Graded Quiz • 30 min

Due Oct 8, 11:59 PM IST



In the above chart, what do the letters in the box plot above represent?

- A = Median, B = Third Quartile, C = First Quartile, D = Inter Quartile Range, E = Minimum, and F = Outliers
- A = Mean, B = Third Quartile, C = First Quartile, D = Inter Quartile Range, E = Minimum, and F = Maximum
- A = Mean, B = Upper Mean Quartile, C = Lower Mean Quartile, D = Inter Quartile Range, E = Minimum, and F = Outliers
- A = Median, B = Third Quartile, C = Mean, D = Inter Quartile Range, E = Lower Quartile, and F = Outliers

Correct

Correct! A = Mean, B = Third Quartile, C = First Quartile, D = Inter Quartile Range, E = Minimum, and F = Outliers.

[Back](#)

## Graded Quiz: Basic and Specialized Visualization Tools

Due Oct 8, 11:59 PM IST

Graded Quiz • 30 min

7. True or False. The correct combination of function and parameter to create a box plot in Matplotlib is Function = plot, and Parameter = kind with value = "box".

1 / 1 point

 False True **Correct**

Correct! Function = plot, and Parameter = kind with value = "box" does create a box in Matplotlib.

8. When creating a histogram in Matplotlib what is the first step?

1 / 1 point

- Import matplotlib as plt and its scripting interface as mpl.
- Import matplotlib as mpl and its scripting interface as plt.
- Import histogram as plt and its scripting interface as mpl.
- Import histogram as mpl and its scripting interface as plt.

**Correct**

Correct! The first step when creating a histogram in matplotlib is to import matplotlib as mpl and its scripting interface as plt.

← Back Graded Quiz: Basic and Specialized Visualization Tools

Due Oct 8, 11:59 PM IST

Graded Quiz • 30 min

9. True or False. The process of creating a scatter plot involves importing Matplotlib to visualize a large set of data.

1 / 1 point

- True
- False

✓ Correct

Correct! A scatter plot is a type of plot that displays values pertaining to typically two variables against each other. The process of creating a scatter plot involves importing Matplotlib to visualize a large set of data.

10. A \_\_\_ is a type of plot where the length of each bar is proportional to the value of the item that it represents.

1 / 1 point

- Histogram chart
- Bar chart
- Line plot
- Figma plot

✓ Correct

Correct! Unlike a histogram, a bar chart, also known as a bar graph, is a type of plot where the length of each bar is proportional to the value of the item that it represents.